Cell Tracking For Quantitative HeLa Cell Cycle Analysis

Hanxiang Hao [1], Xuwen Wu [2], Meichen Liu [2], Anthony P. Reeves [3].

***Abstract*—The purpose of this paper is to develop an algorithm for HeLa cell cycle analysis. The proposed algorithm is composed by cell segmentation and tracking. Our segmentation algorithm includes binarization, nuclei center detection and nuclei boundary delineating, while the tracking algorithm includes feature matching and mitosis occurrence detection by recurrent neural network and ellipse of inertia. Our chosen testing and training datasets are Histone 2B-GFP HeLa cells provided by Mitocheck Consortium. This paper used Jaccard index to measure the segmentation accuracy and TRA method provided by Cell Tracking Challenge for tracking accuracy. Our results, respectably 69.51% and 76.78%, demonstrated the validity of the developed algorithm in investigation and analysis of cancer cell cycle.**

***Index Terms*—HeLa cell, cell cycle analysis, cell tracking, recurrent neural network, ellipse of inertia.**

I. Introduction

TRACKING moving cells in time-lapse video sequences is a challenging task, required for many applications in both scientific and industrial settings. Properly characterizing how cells move as they interact with their surrounding environment is key to understanding the mechanobiology of cell migration. In this paper we will objectively analyze and evaluate the cell segmentation and tracking methods using real time-lapse microscopy videos of labeled cells, along with computer generated video sequences simulating cells moving in realistic environments.

A. Background

Our goal of this paper is to analyze HeLa cell cycle by cell segmentation and cell tracking. The paper uses HeLa cell as model, and the analysis of HeLa cell cycle is achieved by tracing every original HeLa cells and their daughter cells in video frames according to their time scale. Data could be further processed for cell division rate and mitosis occurrence rate, which is crucial for evaluating drug efficacy[1]. The reason of analysis HeLa cell cycle is due to the market needs for antimitotic drugs[2]. HeLa cell is a cell type in an immortal cell line[2]. It's a most commonly used human cell line and essential to cancer tumor development and pharmacodynamics studies. Therefore, HeLa cell is an important model for cancer research. Furthermore, top three cancers using antimitotic drug is breast cancer, non-small cell lung cancer and pancreatic cancer[3]. All of them have either high incidence rates or high death rates[3], and these cancers' treatment is greatly needed. If drug efficacy could be automatically and quickly evaluated by computer analysis, it will be a great improvement by accelerating the FDA approval process of new antimitotic drugs. That's the reason why HeLa cell cycle analysis would be truly beneficial for cancer treatment. Moreover, the example of HaLa cell is shown in figure 1.

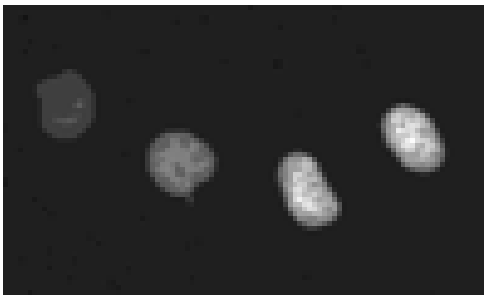

Fig. 1: Example of HeLa Cell.

B. Previous Work

Yang, et al. 2006[4] presents new approaches for cell segmentation and tracking in time-lapse fluorescence image sequence. It first proposes a novel marker-controlled watershed based on mathematical morphology to segment clustered cells with less over-segmentation. Then it employed context information among neighboring frames in order to further segment under-segmented cells or to merge over-segmented cells. After that, it designs a tracking method based on modified mean shift algorithm in which several kernels with adaptive scale, shape, and direction are designed. And then it combined that method with Kalman filter to achieve a better tracking method.

Chen, et al. 2006[5] integrates a series of advanced analysis methods to study the cell cycle process under different conditions of perturbation. It uses the Otsu's algorithm to perform image thresholding, an improved watershed algorithm to resolve the issue of touching nuclei, a postprocessing step is needed to correct and merge the over-segmented nuclei fragments. Then they screen a set of seven features through the set of twelve features to generate the best classification results of different cell stages. And then it uses a matching process to set up the correspondence between nuclei at time t and nuclei at time t+1 by computing the distances between them. Its experimental results show that the method is efficient and effective in cell tracking and phase identification.

[1]Hanxiang Hao is a master student of Cornell University, School of Electrical and Computer Engineering.
E-mail: hh595@cornell.edu
[2]Xuwen Wu and Meichen Liu are master students of Cornell University, School of Biomedical Engineering.
[3]Anthony P. Reeves is the professor of Cornell University, School of Electrical and Computer Engineering, Vision and Image Analysis Group.

Maska, Martin, et al. 2013[6] proposes a tracking scheme involves two steps. First, it applies coherence-enhancing diffusion filtering on each frame to reduce noise and enhance structures. Second, it minimizes the Chan-Vese model in the fast level set-like and graph cut frameworks to detect the cell boundaries. Among the approach, both frameworks have been integrated with a topological prior exploiting the object indication function. The limitations to be addressed in future works are also included. First, a manual separation of cells clustered in the first frame is required which complicates the use of the proposed tracking scheme in experiments. Second, coherence-enhancing diffusion filtering takes up too much time.

Li, Fuhai, et al. 2010[7] presents an automated tracking method for quantitative cell cycle analysis. It introduces a neighboring graph to characterize the spatial distribution of neighboring nuclei, and a novel dissimilarity measure is designed based on the spatial distribution, nuclei morphological appearance, migration, and intensity information in the proposed tracking method. Then, it employs the integer programming and division matching strategy and the novel dissimilarity measure to track cell nuclei.

Li, Shengwen Calvin, et al. 2010[11] defined requirements for clinical applications: sensitivity for single cell detection, real-time positioning, an inducible system, retractable, targeted and durable, monitoring cell fate, and compliant with the FDA GMP guidelines for clinical applications. Methods mentioned in this paper mainly focus on in vivo tracking, and they have their advantages under certain criteria. Improvement on technology provides great benefits for cell tracking (specific cell type, in vivo/non-invasive, real-time, etc.). If combine computation to current cell imaging, we could get better data analyzation results.

Erik Meijering, et al. 2012[13] used computerized image analyzing method for cell tracking, and explained why this method provides benefits. Achieving complete understanding of any living thing inevitably requires thorough analysis of both its anatomic and dynamic properties. Live-cell imaging experiments carried out to this end often produce massive amounts of time-lapse image data containing far more information than can be digested by a human observer. Computerized image analysis offers the potential to take full advantage of available data in an efficient and reproducible manner. A recurring task in many experiments is the tracking of large numbers of cells or particles and the analysis of their (morpho)dynamic behavior.

Cicconet, Marcelo, et al. 2014[9] makes some contributions been in this paper, includes (1) a method for counting embryos in a well, and cropping each individual embryo across frames, to create individual movies for cell tracking; (2) a semi-automated method for cell tracking that works up to the 8-cell stage, along with a software implementation available to the public (this software was used to build the reported database); (3) an algorithm for automatic tracking up to the 4-cell stage, based on histograms of mirror symmetry coefficients captured using wavelets; (4) a cell-tracking database containing 100 annotated examples of mammalian embryos up to the 8-cell stage; and (5) statistical analysis of various timing distributions obtained from those examples.

Chen, Jianxu, et al. 2015[14] focus on cell segmentation and motion tracking in time- lapse images. The known cell tracking approaches mainly fall into two frameworks, detection association and model evolution, each having its own advantages and disadvantages. In experiment, framework was evaluated by 10 different datasets, and author's approach achieved considerable improvement over the state-of-the-art cell tracking algorithms on identifying complete cell trajectories.

C. Paper Outline

We are going to discuss the main algorithm, database and performance metrics for segmentation and tracking in part 2; the results will be included in part 3; and the discussions will be stated in part 4.

II. Methodology and Materials

In this section we will discuss the main algorithm, the performance metrics we used to test our method, and the database we used.

A. Methodology

The proposed algorithm can be divided into two parts: Cell Segmentation and Cell Tracking. The fig. 2 shows a flowchart of our proposed method.

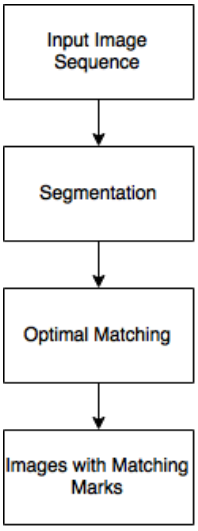

Fig. 2: Flowchart of the proposed algorithm.

In the flowchart, in order to tracking the cells in the original images, we first need to segment these cells. In order to obtain the full shape of each cell, we use watershed technic to extract the cells out of background. Furthermore, after generating the binary segmentation image, we match each cell to the previous frame according to the features of each cell, including *centroid*, *Fourier descriptors* and *Co-occurrence matrix*. During the matching processing, we also optimize our result by detecting the cells in mitosis. The details will be discussed in the following section.

1) Cell Segmentation: Generally, cell segmentation is to detect and localize each cell in each frame of image sequence and then, provides spatial information for the tracking procedure. Essentially, this method consists of three components: binarization, cells center detection and cells boundary delineating.

Normalization:

Since our original image has low contrast and high mean value, we need to normalize the illumination by the following expression:

$$I_{normalized} = \frac{I_{original} - mean(I_{original})}{\sigma(I_{original})} \quad (1)$$

The original image and the normalized image is shown in figure 3.

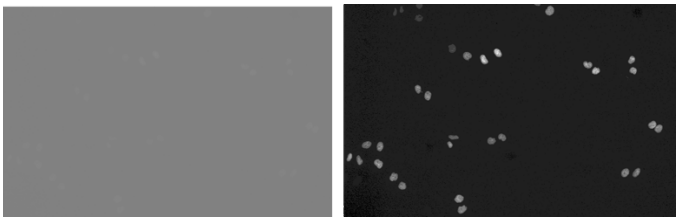

Fig. 3: original image (left) and the normalized image (right).

As the images shown above, after normalization, the cells are extracted from the background.

Binarization:

In order to get a rough binary image with cell-classified pixel as foreground, the proposed method implements adaptive thresholding method in binarization process. The result of adaptive thresholding shows in figure 4.

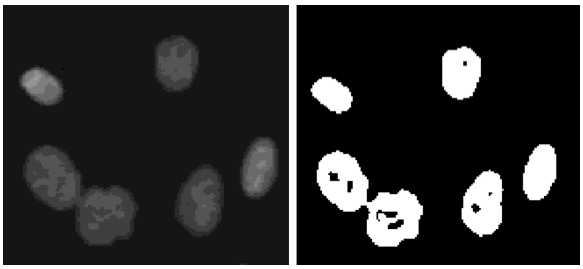

Fig. 4: The normalized image (left) and thresholded image (right).

As shown in the right image, there are some cells containing holes, which is caused by the nuclei of each cell and the structures that have relatively low pixel value. Moreover, since some cells are too close to differentiate, the segmented result classifies them into one single cell. In order to get the full shape for each cell, we need to use watershed technic to deal with this obstacle.

Cells center detection:

In order to use watershed to generate a "good quality" segmented image, we need to find a seed point for each cell. The cells center detection step provides seed points for watershed to find the boundary of each cell. To detect the centers of cells, we use distance map to find the center for each cell. The example of the distance map is shown below.

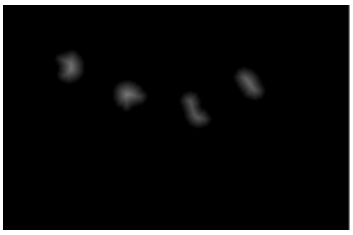

Fig. 5: The example of distance map.

As shown in the figure 5, the center of each cell has relatively larger value than others. In order to obtain the seed points for watershed, we need to find the center (or at least the point inside the cell). Therefore, we use the following approach to obtain the seed points.

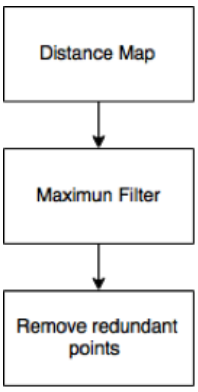

Fig. 6: Generate seed points.

According to the flowchart, the maximum filter removes the peripheral points and keeps the center points. However, there are also some points with the same value in distance map locating in the same cell. We remove these redundant points by limiting the minimum distance for each seed point (the minimum distance usually sets to the same value of maximum filter). The example of the seed points shown below (each white point represents a seed point).

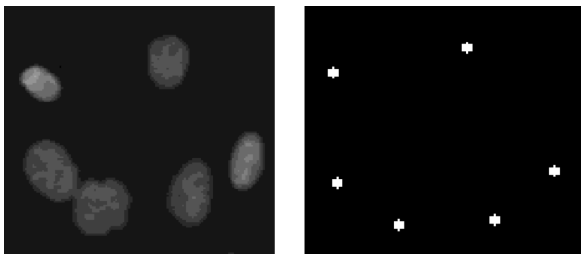

Fig. 7: The normalized image (left) and the seed point for each cell (right).

Cells boundary delineating[5]:

In this procedure, we implement watershed algorithm with the seed points obtained from the previous step. The result of watershed is shown in figure 8.

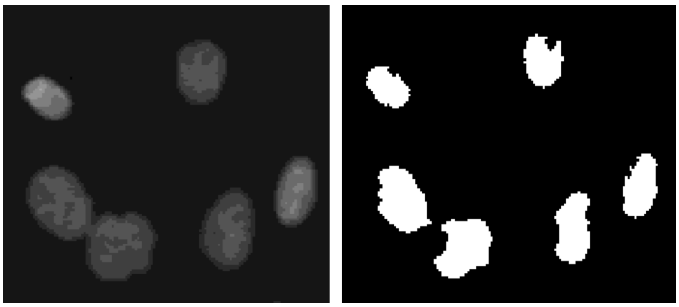

Fig. 8: The normalized image (left) and result of watershed (right).

As the figure below, watershed algorithm can successfully divided the cells overlapping with each other.

Fig. 9: The result of adaptive thresholding (left) and the result of watershed (right).

2) Cells Tracking: In this step, we implement the tracking algorithm with result of the segmented images in the previous step; i.e. matching the segmented cells in two consecutive frames by

- *calculating dissimilarity distance between two cells in two consecutive frames*
- *matching the cells with the smallest dissimilarity distance.*

1) Dissimilarity Measurement.

We first define a feature vector to describe the spatial information for each cell:

$$v_i^k = (c_i^k, f_i^k, m_i^k) \tag{2}$$

where k represents frame number in the image sequence and i represents the label of each cell; $c_i^k = (x_i^k, y_i^k)$ denotes the centroid; f_i^k is a Fourier Descriptors describing the binary shape of a cell, and m_i^k is the texture characteristics measured by co-occurrence matrix for each cell, including Entropy, Energy, Contrast, Homogeneity.

The dissimilarity measure between the ith cell in frame k and the jth cell in frame k+1 is

$$d(v_i^k, v_j^{k+1}) = \alpha_1 * d(c_i^k, c_j^{k+1}) + \alpha_2 * q * d(f_i^k, f_j^{k+1}) + \alpha_3 * q * d(m_i^k, m_j^{k+1}) \tag{3}$$

The three terms $d(c_i^k, c_j^{k+1})$, $d(f_i^k, f_j^{k+1})$ and $d(m_i^k, m_j^{k+1})$ represent the distance measurements for the features we mentioned in the previous section, $\alpha_i, i = 1, 2, 3$ are the weighting parameters which satisfy $\alpha_i \geq 0, \sum_{i=1}^{3} \alpha_i = 1$ and q is the mitosis cell identification parameter, which is determined by Recurrent Neural Network and ellipse of Inertia that we will discuss in the next section. The value of q is defined by:

$$q = \begin{cases} 0, & if\ the\ cell\ is\ in\ mitosis\ process \\ 1, & otherwise \end{cases} \tag{4}$$

The first term $d(c_i^k, c_j^{k+1})$ measures the Euclidean distance between the centroids of two cells, as shown below, where D denotes the maximum migration distance per frame and we set D = 30 pixel empirically[7].

$$d(c_i^k, c_j^{k+1}) = \begin{cases} \frac{\sqrt{(x_i^k - x_j^{k+1})^2 + (y_i^k - y_j^{k+1})^2}}{D}, \\ if\ \sqrt{(x_i^k - x_j^{k+1})^2 + (y_i^k - y_j^{k+1})^2} < D \\ 1, \quad otherwise \end{cases} \tag{5}$$

The second term $d(f_i^k, f_j^{k+1})$ measures the cell shape variation by Fourier Descriptors (FD). Since the lower order of the FD represents the lower frequency component of the figure, we empirically use the order from 1 to 8 to represent the shape of cell. Furthermore, since the Fourier coefficients of 2D figure are a_n, b_n, c_n, d_n, the distance of FD express as:

$$d(f_i^k, f_j^{k+1}) = \sum_{o=1}^{8} \frac{(a_{n,i,o}^k - a_{n,j,o}^{k+1})^2 + (b_{n,i,o}^k - b_{n,j,o}^{k+1})^2 + (c_{n,i,o}^k - c_{n,j,o}^{k+1})^2 + (d_{n,i,o}^k - d_{n,j,o}^{k+1})^2}{max_o(a_{n,i,o}^k - a_{n,j,o}^{k+1})^2 + max_o(b_{n,i,o}^k - b_{n,j,o}^{k+1})^2 + max_o(c_{n,i,o}^k - c_{n,j,o}^{k+1})^2 + max_o(d_{n,i,o}^k - d_{n,j,o}^{k+1})^2} \tag{6}$$

where o represents the order of FD.

The third term $d(m_i^k, m_j^{k+1})$ describes the texture characteristics by co-occurrence matrix for each cell. The features of the co-occurrence matrix are described below:

$$Entropy = -\sum_i \sum_j \overline{P}[i,j] \log \overline{P}[i,j] \tag{7}$$

$$Energy = \sum_i \sum_j (\overline{P}[i,j])^2 \tag{8}$$

$$Contrast = \sum_i \sum_j \overline{P}[i,j](i-j)^2 \tag{9}$$

$$Homogeneity = \sum_i \sum_j \frac{\overline{P}[i,j]}{1-|i-j|} \tag{10}$$

where $\overline{P}$ is the average co-occurrence matrix among the four directions-$0^\circ, 45^\circ, 90^\circ, 135^\circ$-of co-occurrence matrices. Furthermore, the distance of co-occurrence matrix is defined as

$$d(m_i^k, m_i^{k+1}) = \frac{(Ep_i^k - Ep_j^{k+1})^2 + (Eg_i^k - Eg_j^{k+1})^2 + (C_i^k - C_j^{k+1})^2 + (H_i^k - H_j^{k+1})^2}{max(Ep_i^k, Ep_j^{k+1})^2 + max(Eg_i^k, Eg_j^{k+1})^2 + max(C_i^k, C_j^{k+1})^2 + max(H_i^k, H_j^{k+1})^2} \tag{11}$$

where Ep stands for Entropy; Eg stands for Energy; C stands for Contrast and H stands for Homogeneity.

2) Mitosis Identification.

This section describes two methods-ellipse of inertia and Recurrent Neural Network-to identify mitosis occurrence for each cell and in the Result and Discussion section, we will analyze and evaluate the these two methods.

• *Ellipse of Inertia*

This mitosis identification approach is to compute the ellipse of inertia for each cell and calculate the ratio between the major axis and the minor axis in order to differentiate the cell in mitosis or not, since the shape of the cell in mitosis is elongate, and therefore the ratio is larger than the normal cells. The cell that is in mitosis is shown in the figure below.

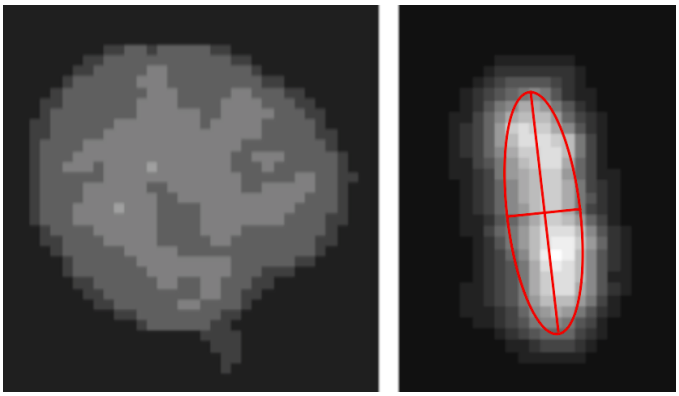

Fig. 10: The cell before mitosis (left) and during mitosis with ellipse of inertia (right).

Obviously, the proposed method is based on the assumption that the cells that are not in mitosis could not be elongated. Although this hypothesis is suitable for most condition, this method will make an incorrect identification when the elongated cell is normal, but not in mitosis.

• *RNN (Recurrent Neural Network)*

A recurrent neural network (RNN) is a class of artificial neural network where connections between units form a directed cycle. This creates an internal state of the network which allows it to exhibit dynamic temporal behavior. Unlike feedforward neural networks, RNNs can use their internal memory to process arbitrary sequences of inputs[15]. The structure of RNN is shown below

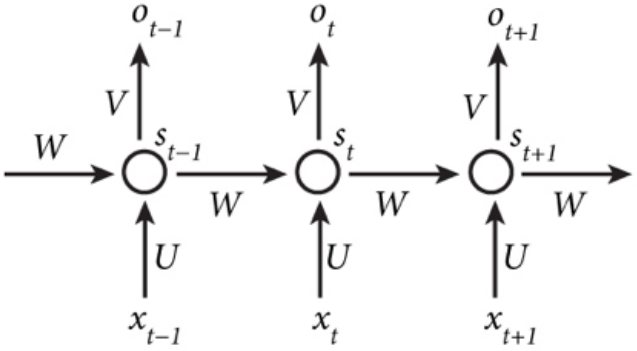

Fig. 11: Structure of RNN [12].

In the case of mitosis identification, we need to make RNN "remember" the cell's shape in the previous frame and therefore predict the state of the cell in the future frame as "alive" or "break". Moreover, the input of RNN x_t is the shape features of each cell in the previous frame and we use the Fourier Descriptors to represent the shape of cells; the output o_t of RNN is the predicted probability of the state ("alive" or "break") for each cell in the next frame.

In our program, the nonlinear function in hidden layer is set as:

$$\tanh(Ux_t + Ws_{t-1}), \tag{12}$$

while the state probability is calculated by softmax which is defined as:

$$o_t = \frac{e^{Vs_t}}{\sum_k e^{Vs_k}} \tag{13}$$

Furthermore, the loss function is defined as:

$$L = -\sum_j y_t \log o_t \tag{14}$$

where y_t is the ground truth of input t. In our program, the input shape feature is the Fourier Descriptors (FD) obtained in the previous step. If the order of FD is set as 10 and the length of hidden layer is set as 100, the shapes of the set of parameters are:

$$x_t \in \mathbb{R}^{40}$$

$$o_t \in \mathbb{R}^{2}$$

$$s_t \in \mathbb{R}^{100}$$

$$U \in \mathbb{R}^{100*40}$$

$$V \in \mathbb{R}^{2*100}$$

$$W \in \mathbb{R}^{100*100}$$

3) Matching Strategy.

The matching strategy associates (matches) the cells in two consecutive frames based on the defined dissimilarity measurement. We compute the dissimilarity for each cell in the two successive frames and find the pair with the smallest distance. If the distance is smaller than our maximum dissimilarity, we match these two cells. Furthermore, in order to match the daughter cells generated after mitosis, we propose an approach for division matching.

Division matching is designed to find sibling cells given the dead cell in the previous frame. To be specific,
1) find all the cells which don't match any cell in the next frame;
2) among these cells, check if there is a sibling cell that satisfies

$$d(f_i^k, f_i^{k+1}) + d(m_i^k, m_i^{k+1}) < d_T \tag{15}$$

where d_T is a maximum dissimilarity of the sibling cell and its parent, which is different with the maximum dissimilarity in the matching strategy;
3) match the parent cell with the two closest sibling cells.

B. Experiment

This section discusses the selection of parameters of the proposed algorithm and the accuracies of segmentation and tracking.

1) Evaluation function: In our paper, we use Jaccard Index to measure the performance of segmentation and use TRA method provided by Cell Tracking Challenge to measure the accuracy of tracking.

The accuracy of the methods (Jaccard), understood as how well the segmented regions of the cells match the actual cell or nuclei boundaries, will be measured by comparing the segmented objects with the ground truth (GT) consisting of the manual annotation of selected frames (2D) and/or image planes (in the 3D cases). Furthermore, the Jaccard similarity index is defined as:

$$Jaccard = \frac{A \cap B}{A \cup B} \tag{16}$$

where A is the set of pixels belonging to the particular object in ground truth and B is the set of pixels belonging to the object in the tested algorithm.

The tracking precision of the methods (TRA) [16] understood as how accurately each given object is identified and followed in successive frames will be based on comparison of acyclic oriented graphs representing the time development of objects in both the GT and each tested method. The nodes of the graph will represent the objects, the edges in the graph and the links between objects in time. Two nodes will be mapped to each other if the objects that they represent have more than 50% overlap. Unmapped nodes in GT mean undetected objects, while unmapped nodes in a computed graph mean false detection or late detection of a split event. Forking in GT and no forking in a computed graph means undetected splitting. All these deviations from the GT graph will be penalized. Numerically, the normalized tracking accuracy measure will be defined as

$$1 - \frac{\min TRA_P, TRA_E}{TRA_E} \tag{17}$$

where TRA_P is the cost of transforming a participant graph into a GT reference and TRA_E is the cost of creating the GT reference from scratch (i.e., it is TRA for empty tracking results). The min operator in the numerator prevents from having final negative results when the participant results are completely wrong. The details of the TRA metrics can be found in the reference [16].

2) Documented Data Set: The database we use is HeLa cancer cells provided by Mitocheck Consortium. The data set contains 2 image sequences with 91 frames, respectively and each image sequence has the ground truth of both segmentation and tracking. All the images format as tif with the size of 700*1100 and the example of original image and its ground truth is shown below.

Fig. 12: The the original image, its segmentation ground truth, and its tracking ground truth. (revise the value for visibleness)

The ground truth for the real data is provided by the challenge, which are all annotated manually. The ground truth for segmentation is given by randomly choosing of certain cells from certain frames. The ground truth for tracking is given to denote the position of the cells, but not with the precise shape.

III. Results

To segment cells from background, we set the maximum distance between each seed point to 20 pixels, since according

to our dataset, the nearest cells have the distance larger than 20 pixels, normally. Furthermore, the Jaccard accuracy of our segmented result is 69.51%.

In order to find an optimized set of parameters for the highest tracking accuracy, we change the three weights in the dissimilarity measurement equation and the maximum dissimilarity between two cells. The figures below shows the results of changing parameters.

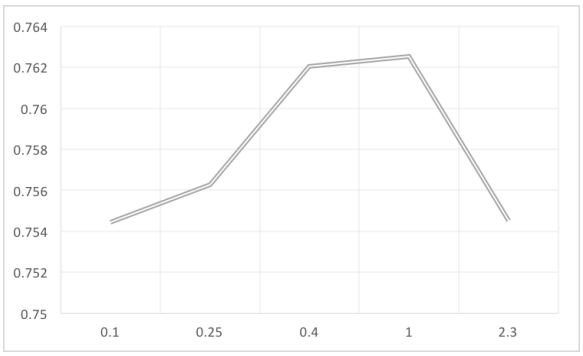

Fig. 13: Changing the ratio of $\frac{\alpha_2}{\alpha_3}$ with fixed $\alpha_1 = 0.5$ and $MaxDissimilarity = 0.4$.

In fig. 13, we change the ratio of $\frac{\alpha_2}{\alpha_3}$ with fixed α_1 and $MaxDissimilarity$, since α_2 and α_3 represent shape and texture, respectively, which are independent with position represented by α_1. Furthermore, the highest accuracy-76.25%- is obtained when the ratio equals to 1.

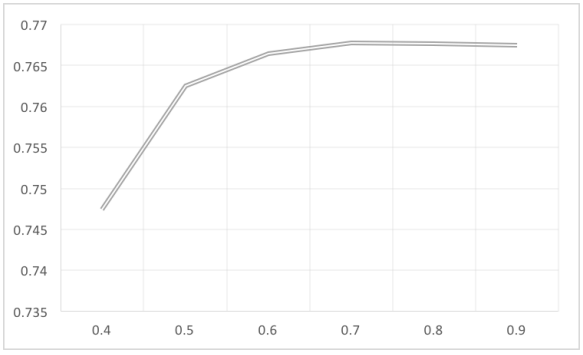

Fig. 14: Changing α_1 with fixed ratio of $\frac{\alpha_2}{\alpha_3} = 1$ and $MaxDissimilarity = 0.4$.

In fig. 14, we change α_1 with fixed ratio of $\frac{\alpha_2}{\alpha_3} = 1$ and $MaxDissimilarity = 0.4$ and obtain the highest accuracy-76.78%-when $\alpha_1 = 0.7$.

In fig. 15, we change max dissimilarity with fixed α_1, α_2, and α_3 and obtain the highest accuracy-76.78%-when the dissimilarity=0.4.

The parameters and the accuracy we got from the experiment are shown below.

$$\alpha_1 = 0.70$$

$$\alpha_2 = 0.15$$

$$\alpha_3 = 0.15$$

$$Max\_Distance = 0.4$$

$$Tracking\ Accuracy = 76.78\%$$

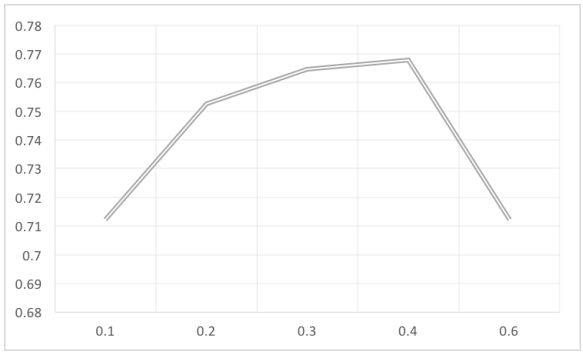

Fig. 15: Changing Max Dissimilarity with fixed $\alpha_1 = 0.7$, $\alpha_2 = 0.15$, and $\alpha_3 = 0.15$.

Furthermore, the table 2 shows the accuracy comparison of the two proposed mitosis identification methods: RNN and ellipse of inertia.

TABLE I: Comparison of Mitosis Identification Methods

RNN	ellipse of Inertia
76.78%	74.82%

IV. Discussion

This section discusses the main findings in the paper, and furthermore, we also give special case that our algorithm cannot deal with.

A. Discussion of the Cell Segmentation

An accuracy segmentation algorithm is significantly important for our algorithm, since if the segmentation method misses too many cells, the cell tracking algorithm cannot have a high accuracy. The result of our segmentation algorithm is shown below.

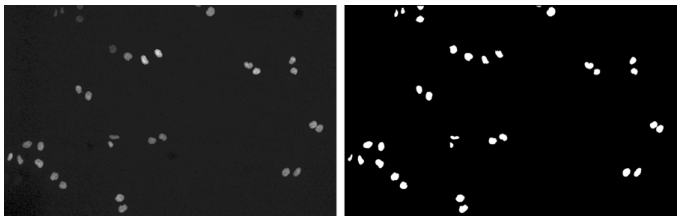

Fig. 16: Normalized image (left) and Segmented image (right).

As mentioned in the section 2, we use watershed algorithm to obtain a full shape for each cell, and therefore, the selection of seed points is extremely important for our algorithm. As shown in the figure 17, our final watershed image misses 2 cells, which is caused by missing seed points as shown in the middle image. Recall that the method to find the seed points is by computing the distance map from the adaptive thresholding image. Since the distance value of the elongate cell is too small in some case, it is possible missing some seed points on our final result. Therefore, in order to improve our segmentation accuracy, we need to implement another useful threshold algorithm, like adaptive multi-level threshold to find seed point for each cell.

Furthermore, our algorithm cannot deal with the under-segmentation problem if two cells are too close to find two

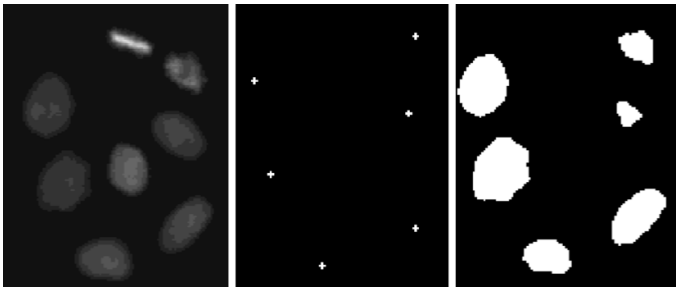

Fig. 17: Example of Missing Cells
(normalized image on the left, seed points on the middle and watershed image on the right).

seed points, individually, as the fig. 18 shown below. In order to segment this piece, we need to use the two split cells in the next frame and resegment it according to the position and shape of the two cells.

Fig. 18: Under-segmentation Error.

B. Discussion of the Cell Tracking

Our tracking algorithm can successfully track the cells in successive frames, as the figure 19 and 20 shown below.

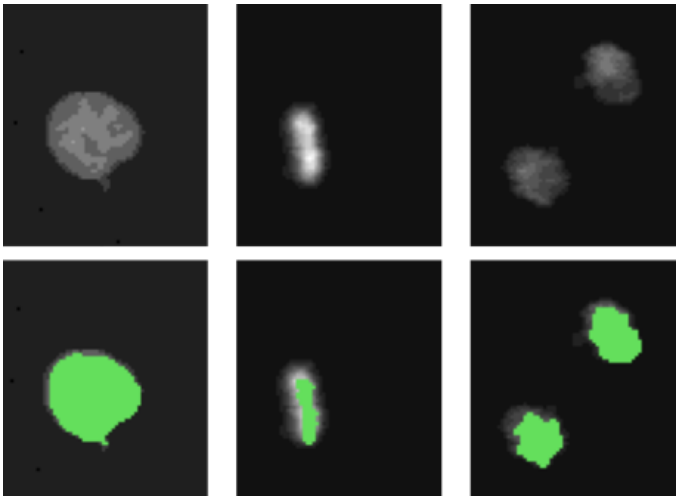

Fig. 19: Example of Cell Tracking Algorithm
Note: The top three images are a cell in three successive frames and the bottom three are our tracking result (the same color represents the cell in the same family tree).

Recall from the section 2, in order to find the siblings after mitosis, we need to find all the cells which are in apoptosis (death) in frame (t+1). However, if the sibling cell doesn't satisfy the condition in the equation [15], we couldn't

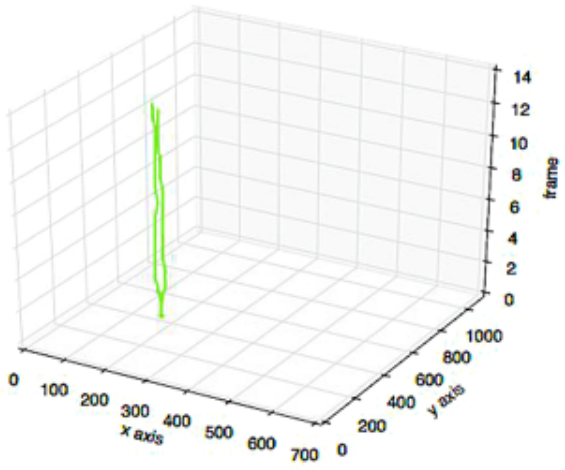

Fig. 20: Result of Cell tracking in successive Frame.

match them even they belong to the same family, as the figure 21 shown.

Fig. 21: Example of False Matching
Note: the same color represents the cell in the same family tree and the number is the unique ID for each cell.

The example in fig. 21 shows a matching error caused by shape dissimilarity between the two cells in the same family.

V. Conclusion

This paper proposed a method to track moving cells in time-lapse video sequences. We first implemented adaptive thresholding and watershed technique to obtain a binary shape for each cell, and then using feature matching by defining a feature vector for each cell, including Centroid, Fourier Descriptors and Co-occurrence Matrix matched cells to previous frame. Moreover, this paper used Jaccard index to measure the segmentation accuracy and TRA method to measure the tracking accuracy. Our results, respectably 69.51% and 76.78%, demonstrated the validity of the developed algorithm in investigation of cancer cell cycle.

References

[1] Browder, T., Butterfield, C. E., Krling, B. M., Shi, B., Marshall, B., O'Reilly, M. S., & Folkman, J. (2000). Antiangiogenic scheduling of chemotherapy improves efficacy against experimental drug-resistant cancer. Cancer research, 60(7), 1878-1886.

[2] 9dresearchgroup. (2015, January). Paclitaxel Market 2015 - Global Industry Size, Trends, Growth, Share, Opportunities and Forecast by 2019.

[3] American Cancer Society. (2016, May 4). Key statistics about breast cancer.

[4] Yang, X., Li, H., & Zhou, X. (2006). Nuclei segmentation using marker-controlled watershed, tracking using mean-shift, and Kalman filter in time-lapse microscopy. Circuits and Systems I: Regular Papers, IEEE Transactions on, 53(11), 2405-2414.

[5] Chen, X., Zhou, X., & Wong, S. T. (2006). Automated segmentation, classification, and tracking of cancer cell nuclei in time-lapse microscopy. Biomedical Engineering, IEEE Transactions on, 53(4), 762-766.
[6] Maka, M., Danek, O., Garasa, S., Rouzaut, A., Munoz-Barrutia, A., & Ortiz-de-Solorzano, C. (2013). Segmentation and Shape Tracking of Whole Fluorescent Cells Based on the Chan?Vese Model. Medical Imaging, IEEE Transactions on, 32(6), 995-1006.
[7] Li, F., Zhou, X., Ma, J., & Wong, S. T. (2010). Multiple nuclei tracking using integer programming for quantitative cancer cell cycle analysis. Medical Imaging, IEEE Transactions on, 29(1), 96-105.
[8] Maka, M., Ulman, V., Svoboda, D., Matula, P., Matula, P., Ederra, C., & Karas, P. (2014). A benchmark for comparison of cell tracking algorithms. Bioinformatics, 30(11), 1609-1617.
[9] Cicconet, M., Gutwein, M., Gunsalus, K. C., & Geiger, D. (2014). Label free cell-tracking and division detection based on 2D time-lapse images for lineage analysis of early embryo development. Computers in biology and medicine, 51, 24-34.
[10] Bensch, R., & Ronneberger, O. (2015, April). Cell segmentation and tracking in phase contrast images using graph cut with asymmetric boundary costs. In Biomedical Imaging (ISBI), 2015 IEEE 12th International Symposium on (pp. 1220-1223). IEEE.
[11] Li, S. C., Tachiki, L. M. L., Luo, J., Dethlefs, B. A., Chen, Z., & Loudon, W. G. (2010). A biological global positioning system: considerations for tracking stem cell behaviors in the whole body. Stem Cell Reviews and Reports, 6(2), 317-333.
[12] http://www.wildml.com/wp-content/uploads/2015/09/rnn.jpg
[13] Meijering, E., Dzyubachyk, O., & Smal, I. (2012). Methods for cell and particle tracking. Methods Enzymol, 504(9), 183-200.
[14] Chen, J., Mahserejian, S., Alber, M., & Chen, D. Z. (2015). A Hybrid Approach for Segmentation and Tracking of Myxococcus Xanthus Swarms. In Medical Image Computing and Computer-Assisted InterventioMICCAI 2015 (pp. 284-291). Springer International Publishing.
[15] Wikipidia: https://en.wikipedia.org/wiki/Recurrent_neural_network
[16] Cell Tracking Challenge 3rd Edition: http://www.codesolorzano.com/celltrackingchallenge/Cell_Tracking_Challenge/Metrics.html